

Team 137: Forecasting Economic Recovery from COVID-19

Members: Bao Pham, Shangshang Song, Sashwat Venkatesh, Kelsey Watkins

Introduction

This study seeks to develop a comprehensive system capable of forecasting economic recovery and identifying cyclical trends following health crises, specifically focusing on the ongoing recovery from the COVID-19 pandemic. The complexity of this problem necessitates a multifaceted approach, incorporating key economic indicators such as GDP and unemployment rates and evaluating the suitability of various forecasting models under diverse conditions.

Understanding economic recovery in the context of COVID-19 is particularly critical given the variability in government responses and policy interventions across nations, which have resulted in markedly different recovery trajectories. This research aims to explore whether a generalized framework for forecasting economic resilience can be established while also acknowledging the inherent diversity in economic structures and recovery dynamics. Thus, it aims to address the challenge of finding an adaptable "best-fit" model for a non-uniform global context.

Problem Definition

The COVID-19 pandemic has presented an unparalleled global challenge, not only in terms of public health but also in its profound economic repercussions. A critical issue that emerged during the pandemic was the difficulty faced by nations in accurately assessing and modeling the trajectories of economic recovery in response to various policy interventions. This investigation seeks to address this challenge by examining potential correlations between key factors—including COVID-19 case counts, vaccination rates, and the stringency of governmental policies—and their impact on economic recovery. By identifying and quantifying these relationships, this research aims to contribute to a deeper understanding of the economic dynamics that follow large-scale health crises, providing insights to inform more effective recovery strategies in future scenarios.

Literature Survey

Current literature on economic recovery from COVID-19 reveals a fragmented approach, where studies often focus on specific indicators or regions. For instance, works such as [7] explore the economic forecasting of advanced and emerging markets but remain limited in scope due to publication timing. Similarly, [11] highlight the challenges of using traditional econometric models to predict recovery following unprecedented events like COVID-19. Studies such as those by [4] discuss the effectiveness of MIDAS models, suggesting potential in nonlinear approaches for capturing economic trends, while [6] and [2] examine the role of fiscal policies in supporting economic resilience. Moreover, [8] and [9] underscore the importance of industry-specific factors, like digital transformation and the differential recovery of small and medium enterprises, which vary widely across regions and sectors.

Our approach seeks to address these gaps by consolidating these fragmented insights into a comprehensive visualization platform. Drawing from [3] theory on economic recovery from disasters, our platform will allow users to easily access and interpret economic health data, making it more accessible to non-experts. By integrating current data with models that account for sector-specific nuances, we aim to deliver a tool that helps anticipate future recovery patterns in response to health crises.

Proposed Method(s) and Experiments

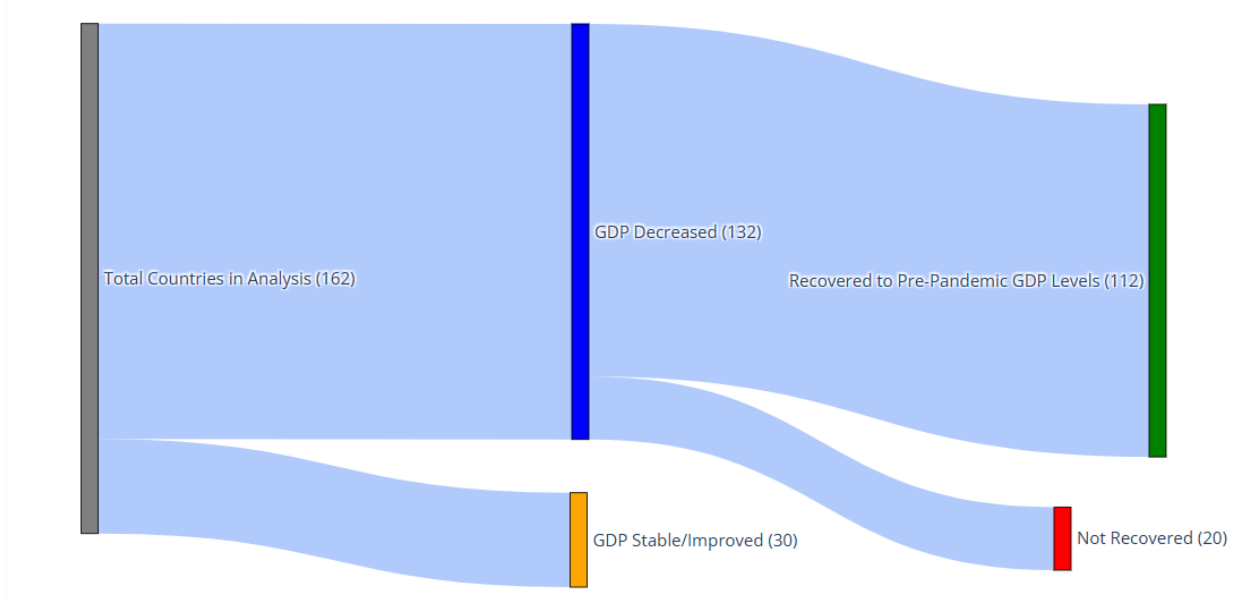
The development of new models to address the unprecedented economic disruptions caused by the COVID-19 pandemic is both intuitive and necessary. As our literature review highlights, traditional approaches primarily relied on linear models to predict economic recovery, which may be insufficient for addressing the unique and complex dynamics of the pandemic. To account for these complexities, our investigation incorporates both linear and non-linear modeling techniques.

For economic metrics, we sourced data from the World Bank Data Bank, obtaining GDP (in 2015 U.S. dollars), GDP per capita (in 2015 U.S. dollars), and population size for 233 entities from 2019 to 2023[13]. Using these data, we calculated the economic impact of the pandemic for each entity by identifying the lowest GDP observed between 2020 and 2022 (“Minimum Pandemic GDP”) and measuring the difference between the Minimum Pandemic GDP and the 2019 GDP. Recovery was quantified by comparing the 2023 GDP to the Minimum Pandemic GDP.

To assess the stringency of public health measures, we utilized the Oxford COVID-19 Government Response Tracker (OxCGRT)[14], which documents 43 different aspects of government responses (e.g., mask mandates, school closures, economic support, vaccine mandates) across 185 entities from January 1, 2020, to December 31, 2022.

The economic and policy datasets were merged using country codes and aggregated at the country level, resulting in a combined dataset covering 162 countries. Among these, 132 countries experienced a Minimum Pandemic GDP below their 2019 GDP. Of these, 112 countries had recovered to either above or below 2019 GDP levels by 2023, while 20 countries had not. Figure 1 (below) illustrates a Sankey diagram illustrating this economic trajectory.

Figure 1. Economic Impact of COVID-19 on Countries



Machine Learning

To investigate the relationships between public health policies and economic recovery, we employed two machine learning models: linear regression and Random Forest regression. These models were chosen to explore both linear and non-linear associations, providing complementary insights into the factors influencing economic recovery.

The response variable, percent GDP recovery, was calculated as the ratio of recovery GDP (2023 GDP) to the GDP decline during the pandemic (2019 GDP minus Minimum Pandemic GDP). This normalization accounts for disparities between countries with higher and lower GDP, enabling fair comparisons across diverse economic contexts. Predictor variables included public health policy measures, economic metrics from 2019, and population size.

For linear regression models, the strength of relationships was assessed using T-test values, while for Random Forest regression, feature importance scores were used to identify key drivers of economic recovery. Given that these models were utilized primarily for explanatory purposes rather than predictive accuracy, model evaluation focused on statistical significance and goodness of fit rather than performance on a hold-out validation set. This approach ensured a robust interpretation of the underlying relationships between policies and recovery outcomes.

Linear Regression

Initially, the analysis included all countries that experienced a decline in GDP during the pandemic years. However, after examining the model residuals, Lithuania was identified as a significant outlier, skewing the results. To improve the model's reliability, Lithuania was removed, and the regression model was refitted. The results are summarized in Figure 2, which highlights the top 10 predictors ranked by the absolute magnitude of their T-values. These predictors include key

public health measures, economic metrics, and policy indicators. Table 1 provides further details about these predictors, including their coefficients, T-values, and p-values.

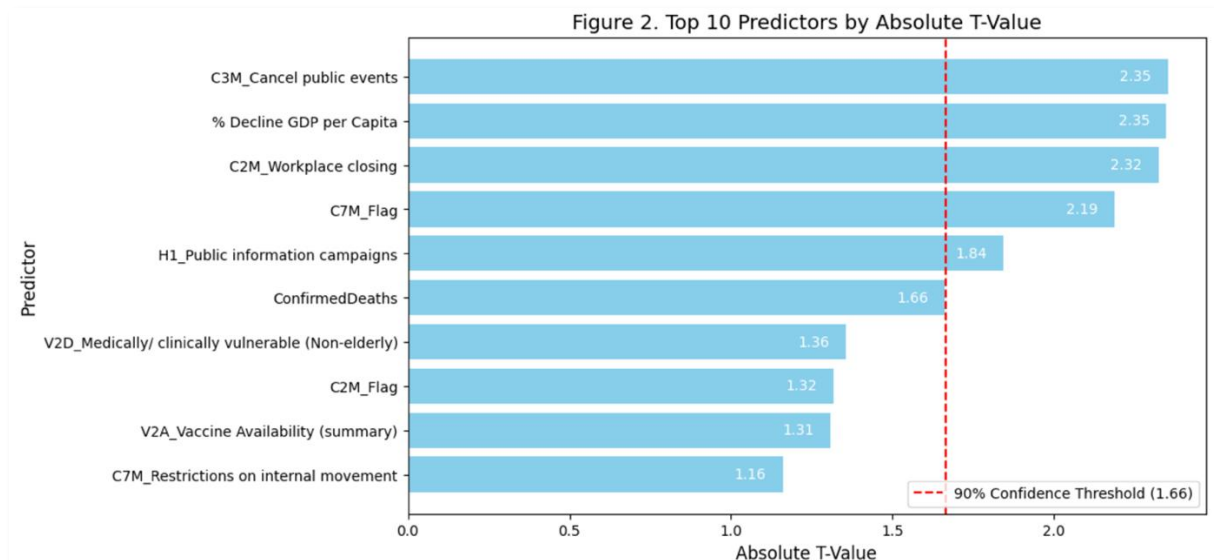


Table 1. Additional Info about Top 10 Predictors

Predictor	Coef	T Value	P-Value
C3M_Cancel public events	1.54E-02	2.366	0.020
% Decline GDP per Capita	-2.83E+01	-2.362	0.021
C2M_Workplace closing	-1.14E-02	-2.337	0.022
C7M_Flag	-1.63E-02	-2.199	0.031
H1_Public information campaigns	1.87E-02	1.855	0.067
Confirmed Deaths	-3.76E-05	-1.670	0.099
V2D_Medically/ clinically vulnerable (Non-elderly)	-1.68E-02	-1.365	0.176
C2M_Flag	1.16E-02	1.325	0.189
V2A_Vaccine Availability (summary)	9.43E-03	1.315	0.192
C7M_Restrictions on internal movement	5.90E-03	1.167	0.247

Random Forest Regression

Given the relatively small sample size (112 rows) compared to the number of predictors (52 variables), overfitting and bias were significant concerns in our analysis. Random Forest, as an ensemble learning method, has an in-built capability for ranking predictors based on their importance. In addition to its feature selection ability, Random Forest is also a non-parametric model and able to capture more complex, non-linear relationships.

Figure 3 illustrates the top 10 predictors identified by Random Forest based on feature importance. Some of the predictor variables ranked most important are the same as in linear regression (% Decline GDP per Capita, C3M_Cancel public events, C2M_Workplace closing). However, Random Forest also ranks other predictors, such as 2019 GDP per Capita as important, while others, such as ConfirmedDeaths, were not considered so.

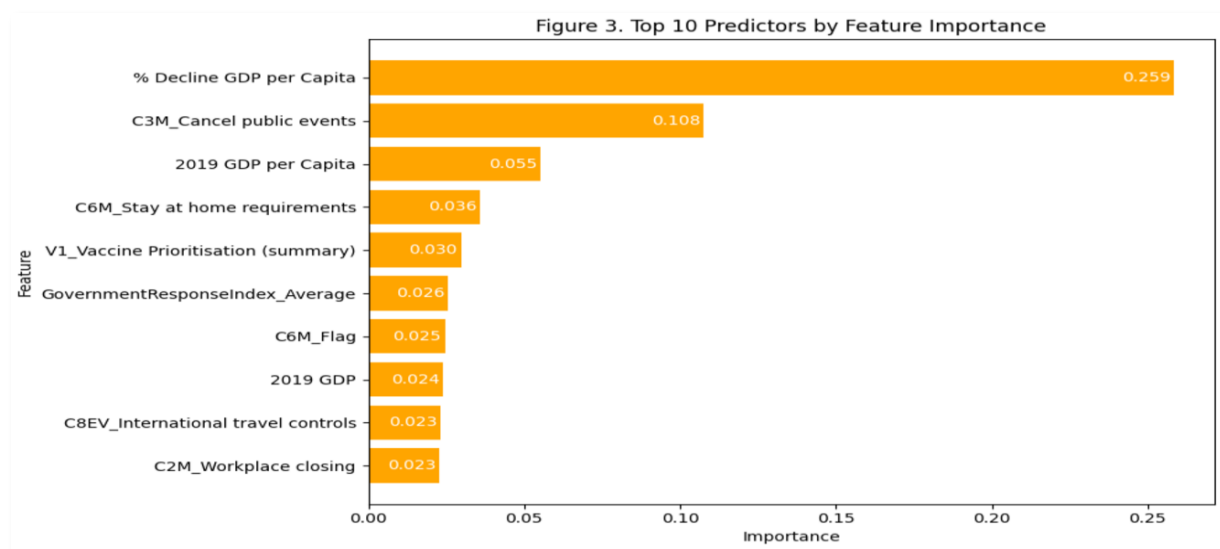


Table 2 shows a summary comparison between the linear regression model and the Random Forest regression model. Since we are using these models for explanatory purposes, we used goodness of fit measures such as R^2 and Adjusted R^2 instead of performance against a holdout set.

Table 2. Summary of Goodness of Fit

Model	R ²	Adjusted R ²
Linear Regression	0.336989	-0.03534
Random Forest Regression	0.883977	0.806628

It seems that Random Forest outperforms linear regression in fitting to the data, which makes intuitive sense since the relationship between public policies and economic impact is likely not

as straightforward as a simple linear relationship. A major advantage of linear regression, however, when compared with Random Forest is its interpretability—it is much easier to quantify the degree to which predictors affect the economic recovery and in what direction, though feature importances can also be helpful in identifying key decision areas.

Conclusions and Discussion

Statement of Distribution of Effort

All team members have contributed a similar amount of effort to this project.

Summary

In conclusion, our findings demonstrate that incorporating non-linear models, such as Random Forest regression, significantly enhances the ability to predict national economic recoveries from public health crises based on policy decisions. These models outperform traditional linear regression approaches by capturing complex, non-linear relationships that better reflect the multifaceted nature of such scenarios. Quantitatively, the Random Forest model's superior goodness-of-fit measures highlight its effectiveness in identifying key predictors of recovery, such as GDP decline, public event cancellations, and workplace closures.

From a real world-impact perspective, our analysis underscores the limitations of traditional linear models in addressing unprecedented global challenges, such as the COVID-19 pandemic. While our study evaluates a limited set of factors, the results indicate that modern crises demand equally modern modeling approaches. Expanding the range of data sources and predictors would allow for a more comprehensive analysis and provide policymakers with actionable insights to inform strategies for recovery and resilience.

Future Directions

There are several promising avenues for future research:

1. **Segmentation and Clustering:** Countries could be segmented using existing classifications (e.g., geographic areas, OBDC status, socioeconomic status) or through unsupervised clustering algorithms. This would enable the identification of region- or segment-specific recovery patterns and policy impacts.
2. **Binary Recovery Analysis:** Transforming recovery into a binary variable (e.g., recovered vs. not recovered) could provide a new perspective. Logistic regression or classification models could be applied to evaluate the impact of predictors on the likelihood of recovery.
3. **Incorporating Additional Predictors:** Future studies could include a wider range of predictors, such as proximity to the origin of the pandemic, population density, literacy rates, healthcare infrastructure, and digital connectivity. These factors could further enrich the analysis and improve the explanatory power of the models.

References

- [1] Ball, L., Leigh, D., & Mishra, P. (2022). Understanding US Inflation during the COVID-19 Era. Brookings Papers on Economic Activity.
<https://dx.doi.org/10.1353/eca.2022.a901276>
- [2] Brada, J., Gajewski, P., & Kutan, A. (2021). Economic resiliency and recovery, lessons from the financial crisis for the COVID-19 pandemic: A regional perspective from Central and Eastern Europe. *International Review of Financial Analysis*, 74.
<https://doi.org/https://doi.org/10.1016/j.irfa.2021.101658>
- [3] Chang, S., & Rose, A. (2012). Towards a Theory of Economic Recovery from Disasters. *International Journal of Mass Emergencies & Disasters*, 30(2).
<https://doi.org/https://doi.org/10.1177/028072701203000202>
- [4] Foroni, C., Marcellino, M., & Stevanovic, D. (2022). Forecasting the Covid-19 recession and recovery: Lessons from the financial crisis. *International Journal of Forecasting*, 38(2). <https://doi.org/https://doi.org/10.1016/j.ijforecast.2020.12.005>
- [5] Fotiadis, A., Polyzos, S., & Huan, T.-C. (2021). The good, the bad and the ugly on COVID-19 tourism recovery. *Annals of Tourism Research*, 87.
<https://doi.org/https://doi.org/10.1016/j.annals.2020.103117>
- [6] Furman, J., Geithner, T., Hubbard, G., & Kearney, M. (2020). Promoting Economic Recovery After COVID-19.
- [7] Garcia, A. R., & Preciado, A. L. J. (2021). COVID-19 and Economics Forecasting on Advanced and Emerging Countries. *EconoQuantum*, 18(1), 22.
<https://doi.org/10.18381/eq.v18i1.7222>
- [8] Holl, A., & Rama, R. (2024). SME digital transformation and the COVID-19 pandemic: a case study of a hard-hit metropolitan area. *Science and Public Policy*.
<https://doi.org/https://doi.org/10.1093/scipol/scae023>
- [9] Meyer, K., Prashantham, S., & Xu, S. (2021). Entrepreneurship and the Post-COVID-19 Recovery in Emerging Economies. *Management and Organization Review*, 17(5).
<https://doi.org/doi:10.1017/mor.2021.49>
- [10] Susskind, D., & Vines, D. (2020). The economics of the COVID-19 pandemic: an assessment. *Oxford Review of Economic Policy*, 36, S1-S13.
<https://doi.org/10.1093/oxrep/graa036>
- [11] Vrontos, I., Galakis, J., Panopoulou, E., & Vrontos, S. (2024). Forecasting GDP growth: The economic impact of COVID-19 pandemic. *Journal of Forecasting*, 43(4).
<https://doi.org/https://doi.org/10.1002/for.3072>
- [12] Xie, W., Rose, A., Li, S., He, J., Li, N., & Ali, T. (2018). Dynamic Economic Resilience and Economic Recovery from Disasters: A Quantitative Assessment. *Risk Analysis: An International Journal*, 38(6). <https://doi.org/https://doi.org/10.1111/risa.12948>

- [13] World Bank, World Development Indicators. (2023). [Data file]. Retrieved from <https://databank.worldbank.org/reports.aspx?source=2&series=NY.GDP.MKTP.CD&country>
- [14] Thomas Hale, Noam Angrist, Rafael Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, Emily Cameron-Blake, Laura Hallas, Saptarshi Majumdar, and Helen Tatlow. (2021). "A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker)." *Nature Human Behaviour*.
<https://doi.org/10.1038/s41562-021-01079-8>